

## 11. Hypothesis Testing

*A statistical hypothesis is some assumption about a quantity statement of a population parameter which may be true or untrue and which is accepted or rejected after testing on the basis of the evidence from a random sample.*

The testing of hypothesis starts with an assumption or guess termed as hypothesis that is made about population parameters. The testing of hypothesis is a process of testing significance of a parameter of the population on the basis of the sample. In testing of hypothesis we compute a statistic.

Hypothesis testing enables us to make probability statements about populations parameters. The hypothesis may not be proved absolutely but in practice it is accepted if it has withstood a critical testing. It is always possible that value of statistics differs from the assumed value. If the difference is too small there is likelihood that guessed or hypothesis value is correct but if difference is too high the hypothesis value might be incorrect.

### Characteristics of Hypothesis

- (a) *A hypothesis must be capable of verification in research work, there must be methods and techniques used for data collection and analysis.*
- (b) *A hypothesis must be related to the existing body of knowledge.*
- © *A hypothesis needs to be precise, simple and specific to be able to develop a good hypothesis.*
- (d) *A hypothesis should be capable of being tested within reasonable time.*
- (e) *Hypothesis should be clear and accurate so, as to draw a consistent conclusion.*
- (f) *A hypothesis should be reliable and consistent with established and known facts.*

#### ***A hypothesis has several functions :***

- (a) *Enhance the objectivity and purpose of a research work;*
- (b) *Provide a research with focus and tells a researcher the specific scope of a research problem to investigate*
- (c) *Help a researcher in prioritizing data collection, hence providing focus on the study; and*
- (d) *Enable the formulation of theory for a researcher to specifically conclude what is true and what is not.*

### Procedure of Testing Hypothesis

- Formulation of Hypothesis
- Determination of a suitable level of significance.
- Determination of suitable test statistics.

- Determination of Critical Region
- Performing the necessary computations
- Drawing conclusions and making Decision.

## Basic Concept of Hypothesis

### 1. Null Hypothesis

A hypothesis stated in the hope of being rejected is called a null hypothesis and it is denoted by  $H_0$ .

Suppose if we are to compare method A with method B about its superiority and if we proceed on the assumption that both method are equally good then this assumption is termed as Null Hypothesis.

Otherwise we think that the method A is Superior or method B is Inferior we assume termed alternative hypothesis.

Null hypothesis  $H_0 \longrightarrow H_0 : \mu = \mu H_0 = 0$

Alternative Hypothesis  $\longrightarrow H_a$

### 2. Alternative Hypothesis

$H_a : \mu \neq \mu H_a$

The alternative hypothesis is that the population mean is not equal to 100. (it may be more or less than 100)

$H_a : \mu > \mu H_a$

The alternative hypothesis is that the population mean is greater than 100.

$H_a : \mu < \mu H_a$

The alternative hypothesis is that the population mean is less than 100.

In choice of Null Hypothesis the following considerations are usually kept in view.



Did You Know ?

*Null Hypothesis represents the hypothesis we are trying to reject and alternative hypothesis represents all other possibilities.*

If the rejection of a certain hypothesis when it is actually true involves great risk. It is taken Null hypothesis because then much probabilities of rejection when it is true

Null Hypothesis should always be specific hypothesis i.e. it should not state approximately a certain value.

(If a hypothesis is type of  $\mu = \mu H_0$  then we call such hypothesis as simple (or Specific) but it is type of  $\mu \neq \mu H_0$  or  $\mu > \mu H_0$  or  $\mu < \mu H_0$  then we call composite or Non Specific Hypothesis)

### 3. The Level of Significance (α)

The significance level, also denoted as alpha or  $\alpha$ , is the probability of rejecting the null hypothesis when it is true. For example significant is 5%. It means out of 100 we would reject about 5 cases. When they should be actually accepted  $(1 - \alpha)$  is probability of accepting a true null hypothesis and it is referred to as confidence level  $(1 - 5\%) = 95\%$ .

The common level of significance and the corresponding confidence level are given below :

- The level of significance 0.10 is related to the 90% confidence level.
- The level of significance 0.05 is related to the 95% confidence level.
- The level of significance 0.01 is related to the 99% confidence level.

The rejection rule is as follows :

- If  $p\text{-value} \leq \text{level of significance } (\alpha)$ , then reject the null hypothesis  $H_0$ .
- If  $p\text{-value} > \text{level of significance } (\alpha)$ , then do not reject the null hypothesis  $H_0$ .

#### 4. Type I and Type II Errors

Type I error means rejection of Hypothesis which should have been accepted and Type II error means accepting Hypothesis which should have been rejected.

Type-I error is denoted by  $\alpha$  (alpha) known as  $\alpha$  error also called the level of significance of test. Type-II error is denoted by  $\beta$  (Beta) known as  $\beta$  error.

Decision		
Accept $H_0$	Reject $H_0$	
Correct Decision ( $1 - \alpha$ )	Type I (Wrong Decision) error ( $\alpha$ error)	$H_0$ (True)
Type II (Wrong Decision) error ( $\beta$ error)	Correct Decision ( $1 - \beta$ )	$H_0$ (False)

In probability :

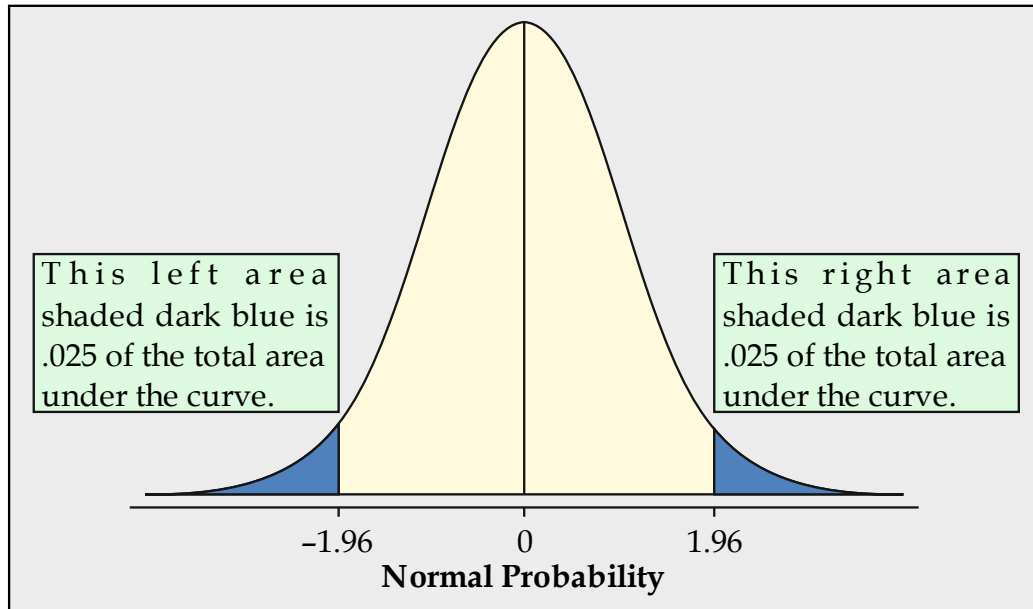
- $\alpha = P(\text{Type-I error}) = P(\text{Reject } H_0 / H_0 \text{ True})$
- $\beta = P(\text{Type-II error}) = P(\text{Accept } H_0 / H_0 \text{ False})$

#### 5. Error in Hypothesis Testing

- If Null Hypothesis ( $H_0$ ) is true but testing is rejected (type I error =  $P(\alpha)$ )
- If Null Hypothesis ( $H_0$ ) is false but testing is accepted (type II error =  $P(\beta)$ )
- If Null Hypothesis ( $H_0$ ) is true but testing is accepted (Right Decision  $P = (1 - \alpha)$ )
- If Null Hypothesis ( $H_0$ ) is false and testing is rejected (Right Decision  $P = (1 - \beta)$ )

#### 6. Two Tailed and One Tailed Test

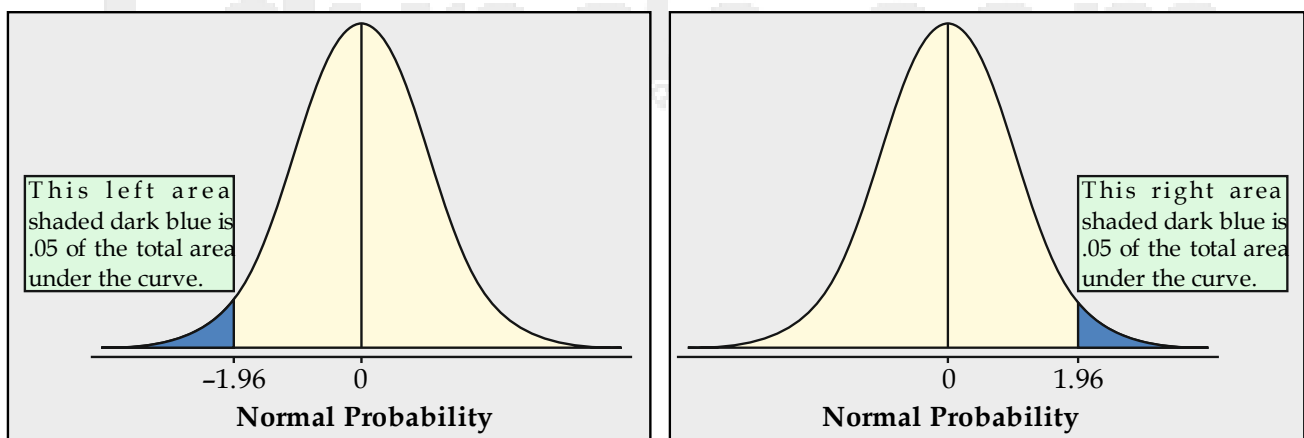
- The two ways of carrying out statistical significance test of a characteristic, drawn from the population, with respect to the test statistic, are a one-tailed test and two-tailed test.
- If the rejection region lies both side of the acceptance region, then the test is called two tailed test.
- If we are using a significance level of 0.05, a two-tailed test allots half of alpha to testing the statistical significance in one direction and half of your alpha to testing statistical significance in the other direction. This means that .025 is in each tail of the distribution of test statistic.



When using a two-tailed test, regardless of the direction of the relationship hypothesized, these are testing for the possibility of the relationship in both directions. For example, we may wish to compare the mean of a sample to a given value  $x$  using a t-test. Our null hypothesis is that the mean is equal to  $x$ . A two-tailed test will test both if the mean is significantly greater than  $x$  and if the mean is significantly less than  $x$ . The mean is considered significantly different from  $x$  if the test statistic is in the top 2.5% or bottom 2.5% of its probability distribution, resulting in a p-value less than 0.05.

**One tailed or One side test** is a test of hypothesis for which the region of rejection is wholly located at one end of the distribution of the test. If we are using a significance level of .05, a one-tailed test allots alpha to testing the statistical significance in the one direction of interest. This means that .05 is in one tail of the distribution of test statistic. When using a one-tailed test, we are testing for the possibility of the relationship in one direction and completely disregarding the possibility of a relationship in the other direction.

Let's return to our example comparing the mean of a sample to a given value  $x$  using a t-test. Our null hypothesis is that the mean is equal to  $x$ . A one-tailed test will test either if the mean is significantly greater than  $x$  or if the mean is significantly less than  $x$ , but not both. Then, depending on the chosen tail, the mean is significantly greater than or less than  $x$  if the test statistic is in the top 5% of its probability distribution or bottom 5% of its probability distribution, resulting in a p-value less than 0.05. The one-tailed test provides more power to detect an effect in one direction by not testing the effect in the other direction.



When the population mean is either lower or higher than some hypothesized value, one tailed test is considered to be appropriate. Where the rejection region is only on the left tail of the curve it is known as left tailed side or where the rejection region is only on the right tail of the curve it is known as right tailed side.

### Comparison Chart

Basis of Comparison	One-Tailed Test	Two-Tailed Test
Meaning	A statistical hypothesis test in which alternative hypothesis has only one end, is known as one tailed test.	A significance test in which alternative hypothesis has two ends, is called two-tailed test.
Hypothesis	Directional	Non-directional
Region of rejection	Either left or right	Both left and right
Determines	If there is a relationship between variables in single direction.	If there is a relationship between variables in either direction.
Result	Greater or less than certain value.	Greater or less than certain range of values.
Sign in alternative hypothesis	$>$ or $<$	$\neq$

### Critical Value

- The testing of statistical hypothesis is done on the basis of the division of the sample space into two mutually exclusive regions. One regions for acceptance (acceptance region) and other region for rejection (rejection region or critical region) of  $H_0$ .
- Critical values for a test of hypothesis depend upon a test statistic, which is specific to the type of test, and the significance level,  $\alpha$ , which defines the sensitivity of the test. A value of  $\alpha = 0.05$  implies that the null hypothesis is rejected 5% of the time when it is in fact true. The choice of  $\alpha$  is somewhat arbitrary, although in practice values of 0.1, 0.05, and 0.01 are common.
- Critical values are essentially cut-off values that define regions where the test statistic is unlikely to lie; for example, a region where the critical value is exceeded with probability  $\alpha$  if the null hypothesis is true. The null hypothesis is rejected if the test statistic lies within this region which is often referred to as the rejection region(s).

### Tests of Hypothesis

Hypothesis testing determines the validity of the assumption (technically described as null hypothesis) with a view to choose between two conflicting hypotheses about the value of a population parameter. It helps to decide on the basis of a sample data, whether a hypothesis about the population is likely to be true or false. There are several tests of hypotheses (also known as the tests of significance) which can be classified as:

- Parametric tests or standard tests of hypotheses; and
- Non-parametric tests or distribution-free test of hypotheses.

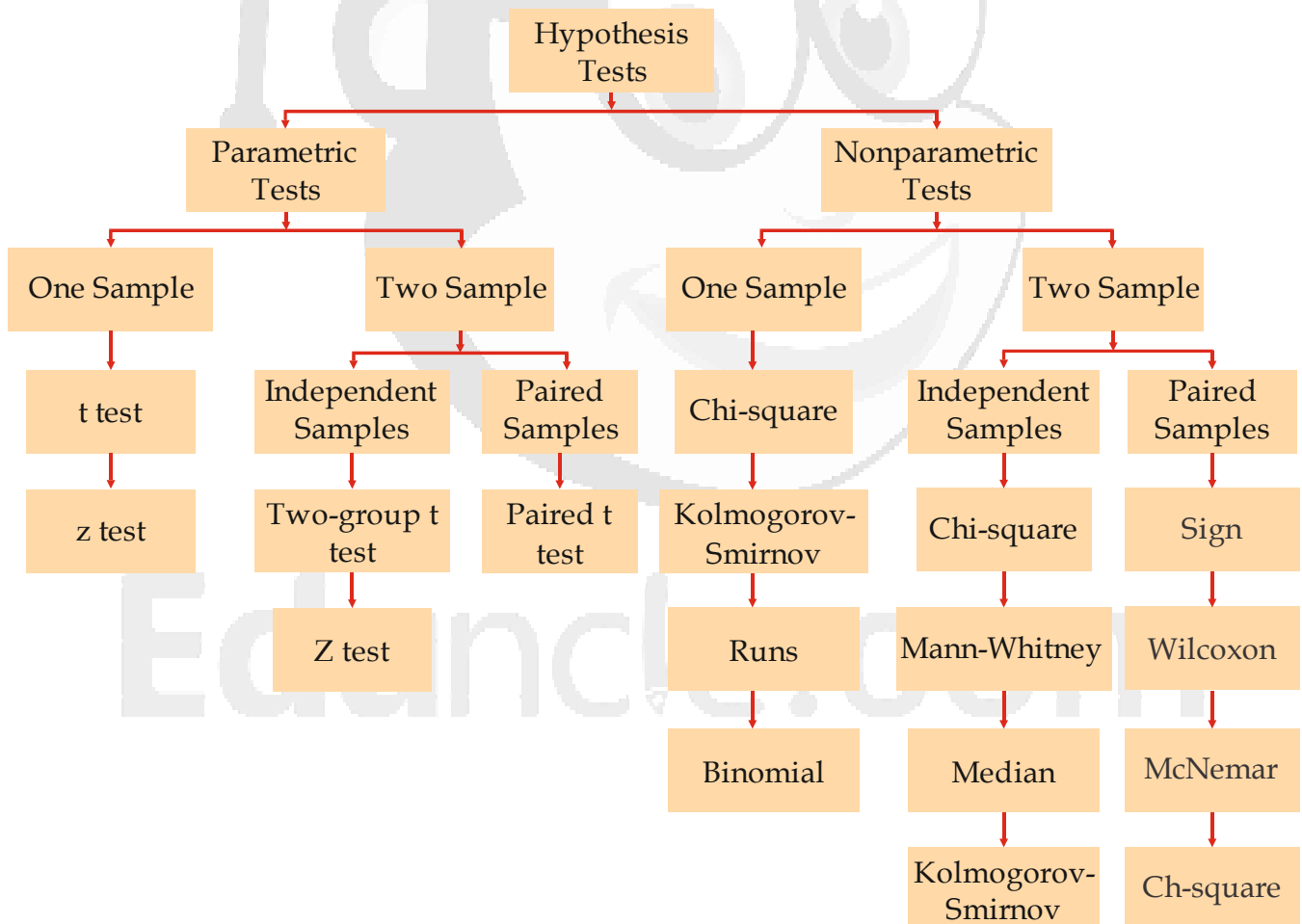
## Parametric Test and Non-Parametric Test

Parametric tests usually assume certain properties of the parent population from which the samples are drawn. Assumptions like observations come from a normal population, sample size is large, assumptions about the population parameters like mean, variance, etc., must hold good before parametric tests can be used, but there are situations when the researcher cannot or does not want to make such assumptions. In such situations researcher can use statistical methods for testing hypotheses which are called non-parametric tests because such tests do not depend on any assumption about the parameters of the parent population. Most non-parametric tests assume only nominal or ordinal data, whereas parametric tests require measurement equivalent to at least an interval scale. As a result, non-parametric tests need more observations than parametric tests to achieve the same size of Type I and Type II errors.

The important parametric tests are:

- (1) z-test;
- (2) t-test;
- (3) F-test
- (4)  $\chi^2$ (chi-square) test and
- (5) Analysis of Variance (ANOVA)

All these tests are based on the assumption of normality i.e., the source of data is considered to be normally distributed.





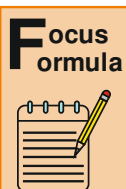
## 1. Z-Test (Large Sample)

Generally if the size of sample exceeds 30, it should be regarded as a large sample. z-test is based on the normal probability distribution and is used for judging the significance of several statistical measures, particularly the mean. The relevant test statistic  $z$ , is worked out and is compared with its probable value which is to be read from the table showing area under normal curve at a specified level of significance for judging the significance of the measure concerned. This test is used even when binomial distribution or t-distribution is applicable on the presumption that such a distribution tends to approximate normal distribution as 'n' becomes larger ( $n > 30$ ). z-test is generally used for:

- (i) Comparing the mean of a sample to some hypothesized mean for the population in case of large sample
- (ii) When population variance is known
- (iii) For judging the significance of difference between means of two independent samples in case of large samples
- (iv) When population variance is known
- (v) z-test is also used for comparing the sample proportion to a theoretical value of population proportion or for judging the difference in proportions of two independent samples when  $n$  happens to be large
- (vi) This test may be used for judging the significance of median, mode, coefficient of correlation and several other measures.

### (a) Testing of Hypothesis about Population Mean

We shall first take the hypothesis testing concerning the population parameter  $\mu$  by considering the two-tailed test.



$$H_0: \mu = \mu_0$$

[ $\mu_0$  is hypothesised value of  $\mu$ .]

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

$$\text{where } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$= \frac{s}{\sqrt{n}}$$

If the calculated value of  $z < -z_{\alpha/2}$  or  $> z_{\alpha/2}$ , the null hypothesis is rejected.

If the calculated value of  $z < -z_{\alpha/2}$  or  $> z_{\alpha/2}$ , the null hypothesis is rejected.

(b) If the hypothesis involves a right-tailed test. For example,

$$H_0: \mu \leq \mu_0 \text{ and } H_1: \mu > \mu_0.$$

For the calculated value  $z > z_{\alpha}$ , the null hypothesis is rejected.

(c) If the hypothesis involves a left-tailed test, i.e.,

$$H_0: \mu \geq \mu_0 \text{ and } H_1: \mu < \mu_0.$$

then for the value  $z < -z_{\alpha}$ , the null hypothesis is rejected.

**Example :** The mean lifetime of a sample of 100 light tubes produced by a company is found to be 1,580 hours with standard deviation of 90 hours. Test the hypothesis that the mean lifetime of the tubes produced by the company is 1,600 hours.

**Solution :** The null hypothesis is that there is no significant difference between the sample mean and hypothesis population mean, i.e.,  $H_0 : \mu = \mu_0$  and  $H_1 : \mu \neq \mu_0$ .

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

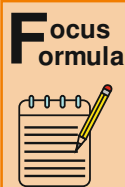
where  $\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$

$$z = \frac{1580 - 1600}{90 / \sqrt{100}} = -2.22$$

The critical is  $z = \pm 1.96$  for a two - tailed test at 5% level of significance. Since, the computed value of  $z = -2.22$  falls in the rejection region, we reject the null hypothesis. Hence, the mean lifetime of the tubes produced by the company may not be 1,600 hours.

### (b) Testing of Hypothesis about The Difference between Two Means

The test statistic for testing the difference between two population means, when the populations are normally distributed.



The z statistic is given by

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

The null hypothesis is  $H_0 : \mu_1 - \mu_2 = 0$

Then, the z statistic is reduced to

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

In case  $\sigma_1^2$  and  $\sigma_2^2$  are not known then for large sample,  $s_1^2$  and  $s_2^2$  can be used instead.

**Example :** You are working as a purchase manager for a company. The following information has been supplied to you by two manufacturers of electric bulbs :

	Company A	Company B
Mean life (in hours)	1300	1,288
Standard deviation (in hours)	82	93
Sample size	100	100

Which brand of bulbs are you going to purchase if you desire to take a risk of 5% ?



**Solution :** Let us take the null hypothesis that there is no significant difference in the quality of the brands of bulbs, i.e.,  $H_0 : \mu_1 = \mu_2$

Since  $\sigma_1^2$  and  $\sigma_2^2$  are not known, therefore, can be replaced by  $s_1^2$  and  $s_2^2$ .

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{1300 - 1288}{\sqrt{\frac{(82)^2}{100} + \frac{(93)^2}{100}}}$$

$$= \frac{12}{\sqrt{67.24 + 86.49}} = \frac{12}{12.399} = 0.968$$

Since our computed value of  $z = 0.968$  is less than critical value of  $z = 1.96$  (5% level), we accept the null hypothesis. Hence, the quality of two brands of bulbs do not differ significantly.

### (c) Test of Hypothesis Concerning Attributes

In case of attributes we can only find out the presence or absence of a certain characteristic. The sampling distribution of the number of successes, being a binomial probability model would have its mean  $\mu = np$  and its standard deviation  $\sigma = \sqrt{npq}$ .

Then 
$$z = \frac{x - np}{\sqrt{npq}} \sim N(0, 1).$$

**Example :** In 600 throws of six-faced die, odd points appeared 360 times, Would you say that the die is fair at 5% level of significance ?

**Solution :** Let us take the hypothesis that die is not biased.

$$p = q = \frac{1}{2}, n = 600, np = 300.$$

Applying the formula;

$$z = \frac{x - np}{\sqrt{npq}} = \frac{360 - 300}{\sqrt{600 \times \frac{1}{2} \times \frac{1}{2}}} = \frac{60}{12.25} = 4.9$$

Since, the computed value of  $z$  is greater than the table value (1.96 at 5% level of significance), the hypothesis is rejected.

Hence, the die does not seem to be fair.

### (d) Testing of Hypothesis about a Population Proportion

The population parameter of interest is population proportion  $\pi$ . If the sample size is large, then sample proportion  $p$  will be approximately normally distributed.

The null hypothesis is that there is no significant difference between the sample proportion and population proportion, i.e.,  $H_0 : p = \pi$

Since the sample proportion  $p$  is unbiased estimator of  $\pi$ ,

**Focus formula**


$$\therefore z = \frac{p - \pi}{\sigma_p};$$

$$\text{where } \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

Therefore, the statistic

$$z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \sim N(0,1)$$

If  $|z| < z_{\alpha}$  the null hypothesis is rejected with  $100\alpha\%$  level of significance.

If  $|z| < z_{\alpha}$ , the null hypothesis is rejected with  $100\alpha\%$  level of significance.

**Example :** A sales clerk in the departmental store claims that 60% of the shoppers entering the store leave without making a purchase. A random sample of 50 shoppers showed that 35 of them left without buying anything. Are these sample results consistent with the claim of the sales clerk ? Use a level of significance of 0.05.

**Solution :** The null hypothesis is

$$H_0 : \pi = 0.60.$$

The sample proportion

$$p = \frac{35}{50} = 0.70.$$

Using the z statistic, we have

$$z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{0.70 - 0.60}{\sqrt{(0.6)(0.4)/50}} = 1.45.$$

The critical value of z is 1.64 at 5% level of significance.

Since, the computed value of  $z = 1.45$  is less than the critical value of  $z = 1.64$ , therefore, the null hypothesis cannot be rejected. Hence based on this sample data, we cannot reject the claim of the sales clerk.

### (e) Testing of Hypothesis about The Difference between Two Proportions

Let  $p_1$  and  $p_2$  be the sample proportions obtained in large samples of sizes  $n_1$  and  $n_2$  drawn from respective populations having proportions  $\pi_1$  and  $\pi_2$ . We can test the null hypothesis that there is no difference between the population proportions, i.e.,

$$H_0 : \pi_1 = \pi_2.$$

The sampling distribution of differences in proportion,  $p_1 - p_2$  is normally distributed with mean

$$\mu_{p_1 - p_2} = \pi_1 - \pi_2$$

and standard deviation

$$\sigma_{p_1-p_2} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$$

Therefore, the statistic is

$$z = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}}$$

If the null hypothesis is true,  $p_1$  and  $p_2$  are two independent unbiased estimators of the same parameter  $\pi_1 = \pi_2 = \pi$ . Thus, our procedure is to pool our observations to obtain the best estimate of the common value  $\pi$ . The pooled estimate of  $\pi$  is the weighted mean of the two sample proportions, i.e.,

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

Our test statistic then becomes

$$z = \frac{p_1 - p_2}{\sigma_{p_1-p_2}},$$

where 
$$\sigma_{p_1-p_2} = \sqrt{p(1-p) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}$$

**Example :** In a random sample of 100 persons taken from village A, 60 are found to be consuming tea. In another sample of 200 persons taken from village B, 100 persons are found to be consuming tea. Do the data reveal significant difference between the two villages so far as the habit of taking tea is concerned ?

**Solution :** Let us take the hypothesis that there is no significant difference between the two villages so far as the habit of taking tea is concerned, i.e.,  $\pi_1 = \pi_2$ .

We are given :

$$p_1 = \frac{x_1}{n_1} = \frac{60}{100} = 0.6, n_1 = 100.$$

$$p_2 = \frac{x_2}{n_2} = \frac{100}{200} = 0.5, n_2 = 200.$$

The appropriate statistics to be used here is given by

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

where 
$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{60 + 100}{100 + 200} = 0.53$$

$$z = \frac{0.6 - 0.5}{\sqrt{(0.53)(0.47)\left(\frac{1}{100} + \frac{1}{200}\right)}} = \frac{0.1}{\sqrt{(0.53)(0.47)(0.015)}}$$

$$= \frac{0.1}{\sqrt{0.0037}} = \frac{0.1}{0.0608} = 1.64.$$

Since, the computed value of  $z$  is less than the critical value of  $z = 1.96$  at 5% level of significance, therefore, we accept the hypothesis. Hence, we conclude that there is no significant difference in the habit of taking tea in the two villages A and B.

## 2. T-Test


A t-test's statistical significance indicates whether or not the difference between two groups' averages most likely reflects a "real" difference in the population from which the groups were sampled. *The t-test is most helpful with a smaller sample size ( $n < 30$ ). The decision depends on the t-statistic and its degrees of freedom (function of sample size).*

A statistically significant t-test result is one in which a difference between two groups is unlikely to have occurred because the sample happened to be a typical. Statistical significance is determined by the size of the difference between the group averages, the sample size, and the standard deviations of the groups. For practical purposes statistical significance suggests that the two larger populations from which we sample are "actually" different.

For applying t-Test in context of small sample the t-value is calculated first of all and then compared with the table value of  $t$  at certain level of significance for given degree of freedom. If the calculated value of  $t$  exceeds the table value (say  $t_{0.05}$  or 5%), we infer that the concerning table value of  $t$  the difference is not treated as significant.

The probability density function of the t-distribution is given by

**Focus  
ormula**



$$f(t) = \frac{Y_0}{\left[1 + \frac{t^2}{v}\right]^{(v+1)/2}} \quad -\infty < t < \infty$$

$Y_0$  is a constant depending on  $n$  such that the total area under the curve is one.

$v = n - 1$  is called the number of degrees of freedom.

**T Test is often called Student's-t test:** *The t-statistic was introduced in 1908 by William Sealy Gosset, a chemist working for the Guinness brewery in Dublin, Ireland. "Student" was his pen name. T test is used to compare two different set of values. It is generally performed on a small set of data.*

### (a) Test of hypothesis about population mean

When the population distribution is normal and standard deviation  $\sigma$  is unknown then the "t" statistic is defined as :

**F**OCUS  
formula


$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

where, sample mean =  $\bar{X}$

hypothesized population mean =  $\mu$ .

Sample size =  $n$

and  $s$  is the standard deviation of the sample calculated by the formula :

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

**Example :** Ten oil tins are taken at random from an automatic filling machine. The mean weight of the tins is 15.8 kg and standard deviation is 0.50 kg. Does the sample mean differ significantly from the intended weight of 16 kg ?

**Solution :** Let the null hypothesis be that the sample mean weight is not different from the intended weight.

Given that

$$n = 10, \bar{x} = 15.8, s = 0.50, \mu = 16$$

Using the t-test, we have

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = \frac{15.8 - 16}{0.50 / \sqrt{10}} = -\frac{0.2}{0.158} = -1.266$$

The table value of  $t$  for 9 d.f. at 5% level of significance is 2.26. The computed value of  $t$  is smaller than the table value of  $t$ . Therefore, the difference is insignificant and the null hypothesis is accepted. Hence the difference between sample mean weight and the intended weight and the intended weight is insignificant.

### (b) Test of Hypothesis about The Difference between Two Means

In testing a hypothesis concerning the difference between the means of two normally distributed populations when the population variances are unknown, the t-test can be used in two types of cases:

**(i) Case of Equal Variances :** Let the null hypothesis be that there is no significant difference between the means of the two populations, i.e.,  $H_0 : \mu_1 = \mu_2$ . When the population variances (though unknown) are equal then the appropriate test statistic to be used is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

will follow t-distribution with  $(n_1 + n_2 - 2)$  d.f., where  $\bar{x}_1$  and  $\bar{x}_2$  are sample means of sample 1 of size  $n_1$  and sample 2 of size  $n_2$  respectively;  $\mu_1$  and  $\mu_2$  are the population means, and  $s$  is "pooled" estimate of the common population standard deviation obtained by pooling the data from both the samples as given below :

### Focus Formula



$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$\text{where } s_1^2 = \frac{\sum (x_1 - \bar{x}_1)^2}{n_1 - 1}$$

$$\text{and } s_2^2 = \frac{\sum (x_2 - \bar{x}_2)^2}{n_2 - 1}$$

(ii) **Case of Unequal Variances** : When the population variances are not equal, i.e.,  $\sigma_1^2 \neq \sigma_2^2$ , we use the unbiased estimators  $s_1^2$  and  $s_2^2$  to replace  $\sigma_1^2$  and  $\sigma_2^2$ . In this case, the difficulty arises because the sampling distribution has large variability than the population variability. The statistic :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}}$$

may not strictly follow t-distribution but may be approximated by t-distribution with a modified value for the degrees of freedom given by

$$\text{d.f.} = \frac{\left[ s_1^2 / n_1 + s_2^2 / n_2 \right]^2}{\frac{(s_1^2 / n_1)^2}{n_1 - 1} + \frac{(s_2^2 / n_2)^2}{n_2 - 1}}$$

**Example** : Two salesman A and B are working in a certain district. From a sample survey conducted by the head Office, the following result were obtained. State whether there is any significant difference in the average sales between the two salesmen.

	A	B
No. of Sales	10	18
Average sales (in lakh Rs.)	190	205
Standard deviation (in lakh Rs.)	20	25

**Solution** : Null hypothesis  $H_0 : \mu_1 = \mu_2$ , i.e., there is no significant difference in the average sales between the two salesmen. Applying t-test

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

$$\text{where } s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$



$$= \sqrt{\frac{9(20)^2 + 17(25)^2}{10 + 18 - 2}} = \sqrt{\frac{3600 + 10625}{26}} = 23.39$$

$$t = \frac{190 - 205}{23.39} \sqrt{\frac{10 \times 18}{10 + 18}} = \frac{15}{23.39} \times 2.54 = 1.63$$

The table value of  $t$  at 5% level of significance for 26 d.f. is 2.056. The calculated value of  $t$  is less than the table value. The hypothesis holds true. Hence, we conclude that there is no significant difference in the average sales between the two salesmen.

### (c) Test of Hypothesis about The Difference between Two Means with Dependent Samples

The samples are dependent if they are paired so that each observation in one sample is associated with some particular observation in the second sample, because of this property, the test we are going to use will be called  $t$ -test. In this test, it is necessary that the observations in the two samples be collected in the form called matched pairs. If two samples are dependent, they must have the same number of units. The appropriate test statistic to be used here is

$$t = \frac{\bar{d}\sqrt{n}}{s}$$

follows  $t$ -distribution with  $(n - 1)$  d.f. where  $\bar{d}$  = mean of the difference is given by  $\bar{d} = \Sigma d/n$ ,  $s$  is the standard deviation of the differences and is given by

$$s = \sqrt{\frac{\Sigma(d - \bar{d})^2}{n - 1}} = \sqrt{\frac{\Sigma d^2 - n(\bar{d})^2}{n - 1}} = \sqrt{\frac{\Sigma d^2}{n - 1} - \frac{(\Sigma d)^2}{n(n - 1)}}$$

and  $n$  is the number of paired observations in the samples. If the computed value of  $t$  (at a specified level of significance for a given number of degree of freedom) is less than the table value of  $t$ , our null hypothesis is accepted, otherwise rejected.

### (d) Test of Hypothesis about Coefficient of Correlation

Testing the hypothesis when the population coefficient of correlation equals zero, i.e.,  $H_0 : \rho = 0$ .

Here the null hypothesis is that there is no correlation in the population, i.e.,  $H_0 : \rho = 0$ . The population coefficient of correlation  $\rho$  measures the degree of relationship between the variables. The appropriate test statistic to be used here is given by :

$$t = \frac{r}{\sqrt{1 - r^2}} \times \sqrt{n - 2}$$

which follow  $t$ -distribution with  $n - 2$  degrees of freedom.

**Example :** In a study of the relationship between expenditure ( $X$ ) and annual sales volume ( $Y$ ), a sample of 10 firms yielded the coefficient of correlation  $r = 0.93$ . Can we conclude on the basis of this data that  $X$  and  $Y$  are linearly related ?

**Solution :** The null hypothesis is  $H_0 : \rho = 0$ , i.e., there is no relationship between two variables. Using the t-test

$$t = \frac{r}{\sqrt{1-r^2}} \times \sqrt{n-2} = \frac{0.93}{\sqrt{1-(0.93)^2}} \sqrt{10-2}$$

$$= \frac{0.93}{\sqrt{0.14}} \times \sqrt{8} = \frac{0.93 \times 2.828}{0.374} = 7.03.$$

The degrees of freedom or  $v = n - 2 = 10 - 2 = 8$ .

The table value of t at 5% level of significance, for 8 d.f. is 2.306. Since the computed value is much greater than the table value, the null hypothesis is rejected. Hence, it may be concluded that X and Y are linearly related.

**Testing the hypothesis when the population coefficient of correlation equals some other value than zero, i.e.,  $H_0 : \rho = \rho_0$ .**

In this case when  $\rho \neq 0$ , the test based on t-distribution will not be appropriate. In testing the hypothesis, the use of Fisher's z-transformation will be applicable. Here, r is transformed into z by

$$z = \frac{1}{2} \log_e \frac{1+r}{1-r}$$

### 3. F-Test

The F-distribution is named in honour of R.A. Fisher who first studied it in 1924. This distribution is usually defined in terms of the ratio of the variances of two normally distributions. The quantity

$$\frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2}$$

is distributed as F-distribution with  $v_1 = n_1 - 1$  and  $v_2 = n_2 - 1$  degrees of freedom, where

$s_1^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n_1 - 1}$  is the unbiased estimator of  $\sigma_1^2$  and  $s_2^2 = \frac{\Sigma(x_2 - \bar{x}_2)^2}{n_2 - 1}$  is the unbiased estimator of  $\sigma_2^2$ .

If  $\sigma_1^2 = \sigma_2^2$ , then the statistic

$$F = \frac{s_1^2}{s_2^2}$$

follows F-distribution with  $n_1 - 1$  and  $n_2 - 1$  degrees of freedom.

The F-distribution sometimes is also called variance ratio distribution which can be seen from the definition. The F-distribution depends on the degrees of freedom  $v_1$  for the numerator and  $v_2$  for the denominator. Therefore, the parameters for F-distribution are  $v_1$  and  $v_2$ . For different values of  $v_1$  and  $v_2$ , we shall have different distributions.

### Test of Hypothesis for Equality of Two Variances

The test of equality of two population variances is based on the variances in two independently selected random sample drawn from two normal populations. Under the null hypothesis  $H_0 : \sigma_1^2 = \sigma_2^2$ .

$$F = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} \text{ reduces to } F = \frac{s_1^2}{s_2^2}.$$

which follows F-distribution with  $v_1$  and  $v_2$  degree of freedom. It is convenient to place larger sample variance in the numerator for computational purpose. If we do so, the ratio of the sample variance will be equal to or greater than one.

If the computed value of F exceeds the table value of F, we reject the null hypothesis, i.e., the alternate hypothesis is accepted.

**Example :** Two sources of raw materials are under consideration by a company. Both sources seem to have similar characteristics but the company is not sure about their respective uniformity. A sample of 10 lots from source A yields a variance of 225 and a sample of 11 lots from source B yields a variance of 200. Is it likely that the variance of source A is significantly greater than the variance of 225 and a sample of 11 lots from source B yields a variance of 200. Is it likely that the variance of source A is significantly greater than the variance of source B at  $\alpha = 0.01$  ?

**Solution :** Null hypothesis is  $H_0 : \sigma_1^2 = \sigma_2^2$ , i.e., the variances of source A and that of source B are same. The F statistic to be used here is

$$F = \frac{s_1^2}{s_2^2}$$

where  $s_1^2 = 225$ , and  $s_2^2 = 200$

$$F = \frac{225}{200} = 1.1$$

The table value of F for  $v_1 = 9$  and  $v_2 = 10$  at 1% level of significance is 4.49. Since the computed value of F is smaller than the table value of F, the null hypothesis is accepted. Hence, the population variances of the two populations are same.

#### 4. Chi-Square Test ( $\chi^2$ )

Chi Square distribution was initially discovered by Hamlet in 1875 but define independently in 1900 by Karl Pearson who gave this notation . This is Non-Parametric Distribution free test since in this case we make no assumption about distribution of parent population.

With the help of  $\chi^2$  test we can know whether a given discrepancy between theory and observation can be attributed to chance or whether it results from the inadequacy of the theory to fit the observed facts. If  $\chi^2$  is zero, it means that the observed and expected frequencies completely coincide. The greater the value of  $\chi^2$ , the greater would be the discrepancy between observed and expected frequency. The formula for computing chi-square is :

**Focus  
Formula**



$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where O = observed frequency

E = expected or theoretical frequency.

The calculated value of  $\chi^2$  is compared with the table value of  $\chi^2$  for given degrees of freedom at specified level of significance. If the calculated value of  $\chi^2$  is greater than the table value, the difference between theory and observation is considered to be significant, i.e., it could not have arisen due to fluctuations of simple sampling. On the other hand, if the calculated value of  $\chi^2$  is less than the table value, the difference between theory and observation is not considered significant, i.e., it could have arisen due to fluctuations of sampling.

**There are two types of chi-square tests. Both use the chi-square statistic and distribution for different purposes :**

- A chi-square goodness of fit test determines if a sample data matches a population.
- A chi-square test for independence compares two variables in a contingency table to see if they are related. In a more general sense, it tests to see whether distributions of categorical variables differ from each another.
- A very small chi square test statistic means that your observed data fits expected data extremely well. In other words, there is a relationship.
- A very large chi square test statistic means that the data does not fit very well. In other words, there isn't a relationship.

#### Assumptions of The Chi-Square Test

- Random sample.
- Independent observations for the sample (one observation per subject).
- All expected counts greater than one.
- No more than 20% of cells with an expected count less than five.

#### Application of Chi-Square

- Testing significance of the sample Variance.
- The goodness of fit of a theoretical distribution.
- The independence of attribute.
- Whether the observed results are consistent with expected results.

#### Test of Goodness of Fit

Tests of goodness of fit are used when we want to determine whether an actual sample distribution matches a known theoretical distribution. *Chi-Square test is known as goodness of fit because it enables us to ascertain how well the theoretical distributions such as Binomial, Poisson, Normal etc. fit empirical distributions, i.e. those obtained from sample data.*

While applying the chi-square test of goodness of fit, the null hypothesis usually states that the sample is drawn from the theoretical population distribution, and the alternative hypothesis usually states that it is not.

**Example :** The number of parts for a particular spare part in a factory was found to vary from day to day. In a sample study, the following information was obtained :

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Total
No. of Parts Demanded	1124	1125	1110	1120	1126	1115	6720

Test the hypothesis that the number of parts demanded does depend on the day of the week.

**Solution :** Let us take the null hypothesis that the number of parts demanded does depend on the day of the week.

The number of spare parts demanded in a week are 6720 and if all days are same, we should expect  $6720/6$ , i.e., 1120 spare parts on a day of the week.

Day	O	E	$(O-E)^2$	$(O-E)^2/E$
Monday	1124	1120	16	0.014
Tuesday	1125	1120	25	0.022
Wednesday	1110	1120	100	0.089
Thursday	1120	1120	0	–
Friday	1126	1120	36	0.032
Saturday	1115	1120	25	0.022
				$\Sigma[(O-E)^2/E] = 0.179$

The table value of  $\chi^2$  for 5 d.f. at 5% level of significance is 11.07. The computed value of  $\chi^2$  is much less than the table value. The null hypothesis is accepted and we conclude that the demand for spare parts is dependent on the day of the week.

**Test of Independence :** One of the most frequent uses of  $\chi^2$  is for testing the null hypothesis that two criteria of classification are Independent. They are independent if the distribution of one criterion in no way depends on the distribution of the other criterion. If they are not independent, there is an association between the two criteria. In the test of independence, the population and sample are classified according to some attributes. The test will indicate only, whether or not any dependency relationship exists between the attributes. It will not indicate the degree of association or the direction of the dependency.

Expected cell frequencies are computed according to the multiplicative rule of probability. If two events are independent, the probability of their joint occurrence is equal to the product of their individual probabilities. Thus, the expected cell frequencies are given by the formula :

$$E_{ij} = \frac{R_i}{N} \times \frac{C_j}{N} \times N = \frac{R_i C_j}{N}$$

To conduct the test, same  $\chi^2$  is employed as discussed earlier, i.e.,

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \text{ or } \sum \frac{(O-E)^2}{E}$$

will follow  $\chi^2$  distribution with  $v = (r - 1)(c - 1)$  degrees of freedom.

While applying the test, the null hypothesis is that the two attributes are independent. If the calculated value of  $\chi^2$  is less than the table value at a specified level of significance, the null hypothesis holds true, i.e., the two attributes are independent. If calculated value of  $\chi^2$  is greater than the table value, the null hypothesis is rejected, i.e., the two attributes are associated.

Chi Square tests-of-independence are widely used to assess relationships between two independent nominal variables. Nominal variable association refers to the statistical relationship(s) on nominal variables. Cross tabulation (also known as contingency or bivariate tables) is commonly used to examine the relationship between nominal variables.

**Example :** A sample of 200 persons with a particular disease was selected. Out of these, 100 were given a drug and the others were not given any drug. The results are as follows :

Number of Persons

	Drug	No Drug	Total
Cured	65	55	120
Not cured	35	45	80
Total	100	100	200

Test, whether the drug is effective or not.

**Solution :** Let us take the null hypothesis that the drug is not effective in curing the disease. Applying  $\chi^2$  test:

The expected cell frequencies are computed as follows :

$$E_{11} = \frac{R_1 C_1}{N} = \frac{120 \times 100}{200} = 60;$$

$$E_{12} = \frac{R_1 C_2}{N} = \frac{120 \times 100}{200} = 60$$

$$E_{21} = \frac{R_2 C_1}{N} = \frac{80 \times 100}{200} = 40;$$

$$E_{22} = \frac{R_2 C_2}{N} = \frac{80 \times 100}{200} = 40$$

The table of expected frequencies is as follows :

60	60	120
40	40	80
100	100	200

O	E	(O-E) <sup>2</sup>	(O-E) <sup>2</sup> /E
65	60	25	0.417
35	40	25	0.625
55	60	25	0.417
45	40	25	0.625
			$\Sigma[(O-E)^2/E] = 2.084$

$$\chi^2 = \Sigma \frac{(O-E)^2}{E} = 2.084$$

$$v = (r - 1) (c - 1) = (2 - 1) (2 - 1) = 1$$

For  $v = 1$ ,  $\chi^2_{0.05} = 3.84$



The calculated value of  $\chi^2$  is less than the table value. The null hypothesis is accepted. Hence, the drug is not effective in curing the disease.

### Yates Correction

*This is an adjustment proposed by yates in calculating the value of chi square for a  $2 \times 2$  contingency table particularly when one or more cell frequencies are less than 5.*

It consists of subtraction of  $\frac{1}{2}$  from the observed frequency of the cell of the first rows and first column and adjusting the cell frequencies of other cell so that row and column totals remain constant.

$$\chi^2(\text{corrected}) = \frac{N \cdot (|ad - bc| - 0.5N)^2}{(a+b)(c+b)(a+c)(b+d)}$$

In case we use the usual formula for calculating the value of chi-square viz.,

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

then Yates' correction can be applied as under :

$$\chi^2(\text{corrected}) = \frac{[|O_1 - E_1| - 0.5]^2}{E_1} + \frac{[|O_2 - E_2| - 0.5]^2}{E_2} + \dots \lim_{x \rightarrow \infty}$$

**Note :** the above formula is used when any of the cell frequency in a  $2 \times 2$  contingency table is less than 5.

### Conditions in Applying Chi-Square Test

- The sample observation should be independent and normally distributed. For the parent population should be greater than 50 or sampling should be replacement.
- While applying chi-square distribution in testing of the goodness of fit or in contingencies table the cell frequencies should not less than 5.
- The data should be expressed in original units rather than percentage or ratio form.
- Sample should be random sample.

*Calculate the chi square statistic  $\chi^2$  by completing the following steps :*

1. *For each observed number in the table subtract the corresponding expected number  $(O - E)$ .*
2. *Square the difference  $[(O - E)^2]$ .*
3. *Divide the squares obtained for each cell in the table by the expected number for that cell  $[(O - E)^2 / E]$ .*
4. *Sum all the values for  $(O - E)^2 / E$ . This is the chi square statistic.*

## 5. Analysis of Variance (ANOVA)

ANOVA is used to test for the significance of the difference between more than two sample means and to make inferences about whether the samples are drawn from the populations having the same mean.

### Assumptions in Analysis of Variance (ANOVA)

1. Each sample is drawn randomly from a normal population and the sample statistics tend to reflect the characteristics of the population.
2. The populations from which the samples are drawn have identical means and variances, i.e.,

$$\mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$$

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_n^2$$

In case we are not in a position to make these assumptions in a particular problem, the analysis of variance technique should not be used.

### Computation of Analysis of Variance

The null hypothesis taken while applying analysis of variance technique is that the means of different sample do not differ significantly.

1. One-way classification, and
2. Two-way classification.

#### 1. One-Way Classification

The term 'one-factor analysis of variance' refers to the fact that a single variable or factor of interest is controlled and its effect on the elementary units is observed. In one-way classification, the data are classified according to only one criterion. The one-way analysis of variance is designed to test the null hypothesis :

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

i.e., the arithmetic means of the population from which the k samples are randomly drawn are equal to one another. The steps involved in carrying out the analysis are :

1. **Calculate the Variance between The Samples :** Sum of squares is a measure of variation. The sum of squares between samples is denoted by SSB. For calculating variance between sample, we take the total of the square of the variations of the means of various samples from the grand mean and divide this total by the degree of freedom. The steps in calculating variance between samples will be :

- (a) Calculate the mean of each sample, i.e.,  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k$
- (b) Calculate the grand mean  $\bar{\bar{X}}$ . Its value is obtained as follows :

$$\bar{\bar{X}} = \frac{\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_k}{n_1 + n_2 + \dots + n_k}$$

- (c) Take the difference between the means of the various samples and the grand mean ;
- (d) Square these deviations and obtain the total which will give sum of squares between the sample; and

- (e) Divide the total obtained in step (d) by the degrees of freedom. The degrees of freedom will be one less than the number of samples, i.e., if there are 4 sample then the degree of freedom will be  $4 - 1 = 3$  or in general  $v = k - 1$ , where  $k$  = number of samples.
2. **Calculate The Variance within The Samples :** The variance (sum of squares) within samples measures those inter-sample differences that arise due to chance only. It is denoted by SSW. For calculating the variance within the samples, we take the total of the sum of squares of the deviation of various items from the mean values of the respective samples and divide this total by the degrees of freedom. Thus, the steps in calculating variance within the samples will be :
- Calculate the mean value of each sample, i.e.,  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k$ ,
  - Take the deviations of the various observations in a sample from the mean values of the respective samples.
  - Square these deviations and obtain the total which gives the sum of squares within the samples.
  - Divide this total obtained in step (c) by the degrees of freedom, the degrees of freedom is obtained by deducting from the total number of observations, the number of sample, i.e.,  $v = n - k$ , where  $k$  refers to the number of samples and  $n$  refers to the total number of all the observations.
3. **Calculate The F-Ratio :** Calculate the F-ratio as follow :

$$F^* = \frac{\text{Variance between the samples}}{\text{Variance within the sample}} \text{ or } F = \frac{s_1^2}{s_2^2}$$

F is always computed with the variance between the sample means as the numerator and the variance within the sample means as the denominator. The denominator is computed by combining the variance within the  $k$  samples into single measures.

4. **Compare The Calculated Value of F :** Compare the calculated value of F with the table value of F for the given degrees of freedom at a certain critical level (generally we take 5% level of significance). If the calculated value of F is greater than the table value of F, it indicates that the difference in sample means is significant i.e., it could not have arisen due to fluctuations of random sampling or, in other words, the samples do not come from the same population. On the other hand, if the calculated value of F is less than the table value, the difference is not significant and hence could have arisen due to fluctuations of random sampling.

## 2. Two-Way Classification

There are, many situations in which the response variable of interest may be affected by more than one factor.

*When it is believed that two independent factors might have an effect on the response variable of interest, it is possible to design the test so that an analysis of variance can be used to test for the effects of the two factors simultaneously. Such a test is called two-factor (way) analysis of variance.*

Thus, with the two-factor analysis of variance, we can test two sets of hypothesis with the same data at the same time.

**F**ocus  
formula


*Anova Table : Two-Way Classification*

Source of Variation	Sum of Squares	d.f.	Mean Square
Between columns	SSC	$c - 1$	$MSC = SSC/(c - 1)$
Between rows	SSR	$r - 1$	$MSR = SSR/(r - 1)$
Residual	SSE	$(c - 1)(r - 1)$	$MSE = SSE/(c - 1)(r - 1)$
Total	SST	$rc - 1$	

SSC = Sum of squares between columns

SSR = Sum of squares between rows

SSE = Sum of squares for the residual

SST = Total sum of squares.

The sum of squares for the source "Residual"\* is obtained by subtracting from the total sum of squares, the sum of squares between columns and rows.

The total number of degrees of freedom =  $cr - 1$

where, c refers to columns and r refers to rows.

Number of degrees of freedom between columns =  $(c - 1)$

Number of degrees of freedom between rows =  $(r - 1)$

Number of degrees of freedom for residual =  $(c - 1)(r - 1)$

The total sum of squares, sum of squares between columns and sum of squares between rows are obtained in the same way as before.

Residual = Total sum of squares - Sum of squares between columns - Sum of squares between rows.

### Mann Whitney Test (U Test)

This test is an analogous of *t*-test for two independent samples. Here we test  $H_0: M_1 = M_2$  against  $H_1: M_1 > M_2$ , or  $H_1: M_1 < M_2$ , or  $H_1: M_1 \neq M_2$  on the basis of two independent samples drawn from continuous populations.

To perform this test, we first of all rank the data jointly. Taking them as belonging to a single sample in either an increasing or decreasing order of magnitude. We usually adopt low to high ranking process which means we assign rank 1 to an item with lowest value, rank 2 to the next higher item and so on. In case there are ties, then we would assign each of the tied observation the mean of the ranks which they jointly occupy. For example, if sixth, seventh and eighth values are identical, we would assign each the rank  $(6 + 7 + 8)/3 = 7$ . After this we find the sum of the ranks assigned to the values of the first sample (and call it  $R_1$ ) and also the sum of the ranks assigned to the values of the second sample (and call it  $R_2$ ). Then we work out the test statistic i.e., U, which is a measurement of the difference between the ranked observations of the two samples as under:

$$U = n_1 \cdot n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

where  $n_1$ , and  $n_2$ , are the sample sizes and  $R_1$  is the sum of ranks assigned to the values of the first sample. (In practice, whichever rank sum can be conveniently obtained can be taken as  $R_1$ , since it is immaterial which sample is called the first sample.)

In applying U-test we take the null hypothesis that the two samples come from identical populations. If this hypothesis is true. It seems reasonable to suppose that the means of the ranks assigned to the values of the two samples should be more or less the same. Under the alternative hypothesis, the means of the two populations are not equal and if this is so, then most of the smaller ranks will go to the values of one sample while most of the higher ranks will go to those of the other sample.

If the null hypothesis that the  $n_1 + n_2$  observations came from identical populations is true, the said 'U' statistic has a sampling distribution with

$$\text{Mean} = \mu_U = \frac{n_1 \cdot n_2}{2}$$

and Standard deviation (or the standard error)

$$= \sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

If  $n_1$  and  $n_2$  are sufficiently large (i.e., both greater than 8), the sampling distribution of U can be approximated closely with normal distribution and the limits of the acceptance region can be determined in the usual way at a given level of significance, but if either  $n_1$  or  $n_2$  is so small that the normal curve approximation to the sampling distribution of U cannot be used.

### Kruskal-Wallis Test (H-test)

This test is conducted in a way similar to the U test described above. *This test is used to test the null hypothesis that 'k' independent random samples come from identical universes against the alternative hypothesis that the medians of these universes are not equal.* This test is analogous to the one-way analysis of variance, but unlike the latter it does not require the assumption that the samples come from approximately normal populations or the universes having the same standard deviation.

In this test, like the U test, the data are ranked jointly from low to high or high to low as if they constituted a single sample. The test statistic is H for this test which is worked out as under:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(n+1)$$

where,  $n = n_1 + n_2 + \dots + n_k$  and  $R_i$  being the sum of the ranks assigned to  $n_i$  observations in the  $i$ th sample.

If the null hypothesis is true that there is no difference between sample medians and each sample has at least five items, then the sampling distribution of H can be approximated with a chi-square distribution with  $(k - 1)$  degrees of freedom. As such we can reject the null hypothesis at a given level of significance if H value calculated, as stated above, exceeds the concerned table value of chi-square.

**Ques.** The power of the statistical hypothesis testing is denoted by :

(NTA UGC-NET Dec. 2015 P-III)

- |                      |                    |
|----------------------|--------------------|
| (1) $\alpha$ (alpha) | (2) $\beta$ (beta) |
| (3) $1 - \alpha$     | (4) $1 - \beta$    |

**Ans.** (4) The power of the statistical hypothesis testing is denoted by  $1 - \beta$ .



**Ques.** Under which of the following situations, chi-square test is applicable ?

- (a) testing homogeneity
- (b) testing goodness of fit
- (c) testing equality of two sample means
- (d) testing equality of two sample proportions
- (e) testing independence of attributes

**Codes :**

(NTA UGC-NET June 2015 P-III)

- |                           |                           |
|---------------------------|---------------------------|
| (A) Only (a), (b) and (c) | (B) Only (a), (b) and (e) |
| (C) Only (c), (d) and (e) | (D) Only (a), (c) and (e) |

**Ans.** (B) Chi-square test is applicable under given below situations :

- testing homogeneity
- testing goodness of fit
- testing independence of attributes

**Ques. Statement-I :** When the null hypothesis is true but as per the hypothesis-testing, it is rejected, it is known as beta type error in hypothesis testing.

**Statement-II :** Chi-square test is exclusively a non-parametric test.

**Codes :**

(NTA UGC-NET July 2016 P-II)

- (1) Both the statements are true
- (2) Both the statements are false
- (3) Statement-I is true while Statement-II is false
- (4) Statement-I is false while Statement-II is true

**Ans.** (2) Both the statements are false.

Eduncle.com